Increased Productivity for Target Compound Screening Using TOF-MS and Accurate-Mass Databases with Optional Retention Time

Part 2: Integrated, Automated Workflows

Technical Overview

Authors

Frank Kuhlmann, Pete Stone, and Ken Imatani
Agilent Technologies, Inc.
Santa Clara, CA USA

Abstract

Agilent provides accurate-mass databases that can be combined with its time-of-flight (TOF) mass spectrometry (MS) systems to perform faster, more reliable analyses for target compounds. Agilent MassHunter Qualitative Analysis software enables highly integrated and automated workflows to search these databases. The software includes sophisticated compound detection algorithms that save hours of analyst time, and allows fully automated analyses from sample injection to final report. These features enable high-throughput screening without compromising data quality, reducing labor costs and ensuring confident results.

Agilent Technologies

Overview

Time-of-flight MS provides accurate-mass spectra that give increased specificity for target compound screening. With high resolution and ppm mass accuracies, the Agilent TOF and quadrupole-TOF (Q-TOF) LC/MS systems greatly reduce the chance of reporting false positive results. To make it much easier to perform screening, Agilent provides separate accurate-mass databases for many compound classes: endogenous metabolites (and related compounds), pesticides, and forensics/toxicology.

As described in part 1 of this technical overview,[1] these Personal Compound Databases (PCDs) and associated software deliver a high degree of specificity and confidence, due to unique scoring algorithms and the ability to add LC retention times to the database when using standardized chromatographic conditions.

Accurate-mass databases are used in two general application areas:

- Target and non-target screening in areas such as forensics and toxicology, food, and environmental analysis.
- Differential analysis of samples or sample sets with the goal of identifying differences. Differential analysis is used frequently in metabolomics, proteomics, and other areas.

This technical overview demonstrates how the Agilent MassHunter Workstation software benefits these type of analyses by providing:

- A dramatic increase in productivity using highly automated compound data processing.
- Flexibility to balance throughput versus sensitivity when screening for large numbers of compounds
- The ability to analyze target *and* unknown compounds in a single analysis

Using databases with accurate mass and optional retention time

The MassHunter family of software products is a tightly integrated suite that enables users to process data very efficiently. The software offers multiple ways to interrogate data and search AMRT databases, so the analyst can choose the workflow that delivers answers most quickly.

Definitions

- Accurate-mass database – **AM database** – contains the calculated exact masses, molecular formulas, and names of compounds, but no spectra
- Accurate-mass/retention time database – **AMRT database** – same as AM database, plus, optionally, contains retention times from a defined set of chromatographic conditions
- **Library** – same as AM database, but contains MS/MS spectra and optional RT
- **Personal Compound Database (PCD)** – AM or AMRT database in MassHunter format, located on the local PC or intranet, provided by Agilent or created by a user
- **Personal Compound Database and Library (PCDL)** – same as PCD, but in addition contains MS/MS spectra for some or all compounds
- **Personal Compound Database and Library (PCDL) software** – MassHunter software module that allows conducting manual searches, adding MS/MS spectra to a library, and editing of AMRT information and MS/MS spectra
- **Target compounds** – compounds being sought
- **Non-target compounds** – compounds not being sought but in database
- **Unknown compounds** – compounds not being sought and not in database

Multiple ways to conduct searches

As one example of the available flexibility, scientists can search AM and AMRT databases from four different places in the MassHunter software:

- Personal Compound Database and Library (PCDL) software
- MassHunter Qualitative Analysis
- Mass Profiler
- Mass Profiler Professional

Each way of searching has unique advantages, as described below.

Manual searches within the Personal Compound Database and Library (PCDL) software

The user interface of the PCDL software provides separate tabs to enable manual search of a single accurate mass or batch search of a list of accurate masses. Users can paste mass lists or import them in .csv or .txt format. The mass lists are typically generated manually from mass spectra or automatically via the Find Compounds algorithms in the Agilent MassHunter Qualitative Analysis software.

The PCDL software includes a Batch Summary tab that allows analysts to make decisions when conflicts appear in the results. For example, if multiple masses return the same hit, the user can assign a different compound to each mass. Likewise, if a single mass returns multiple hits from the database, the user can choose the appropriate compound. In both cases, additional information is necessary in order to resolve those conflicts. Here, adding LC retention time into the search is extremely useful for differentiating compounds.

The batch search mode is useful for updating LC retention times in the database. To do this, the user analyzes a standard mixture, searches the compounds against an AMRT database, and resolves any conflicts. Then the investigator can update all retention times in the database with a single click of the "Apply retention times" button.

Note that the new PCDL software not only provides the ability to edit and search AMRT databases, but also to add MS/MS spectra to the compound entries, which allows users to conduct MS/MS library searches. This additional MS/MS library search capability is not covered in detail in this technical overview.

An Agilent technical note describes in more detail how to conduct searches from the original PCD software, which did not have the MS/MS library search capability.[2]

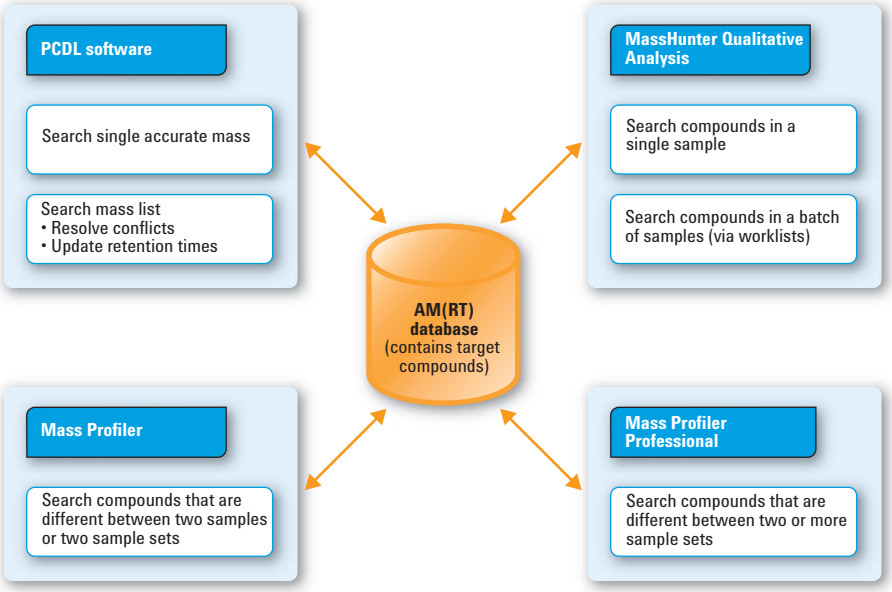
PCDL software
Search single accurate mass
Search mass list
• Resolve conflicts
• Update retention times
MassHunter Qualitative Analysis
Search compounds in a single sample
Search compounds in a batch of samples (via worklists)
AM(RT) database
(contains target compounds)
Mass Profiler
Search compounds that are different between two samples or two sample sets
Mass Profiler Professional
Search compounds that are different between two or more sample sets

Figure 1. For different workflows, users can initiate searches of AM(RT) databases from multiple MassHunter software products.

Searches from within MassHunter Qualitative Analysis software

The most common way to search AM(RT) databases is to initiate the search from MassHunter Qualitative Analysis on all compounds found in a single sample. This type of search can be used to screen for both target and non-target compounds. The software displays the results in MassHunter Qualitative Analysis, where each row in the compound results table lists the database search result with highest score. A separate window for database search results provides more details, listing the best hits with their scores. (The scores are based on accurate mass, isotope ratios, and isotope spacing.) The best hit has a check mark, but the user can override that result and select another hit as the best one.

The entire process can be fully automated so that after each acquisition, the data processing (including the search) is scheduled automatically according to the tasks and parameters defined in the data processing method. Analysts can set up a worklist to automate the data acquisition and processing for an entire batch of samples.

Searches from within Mass Profiler or Mass Profiler Professional

Investigators who perform differential analysis can search AM(RT) databases on compounds that are found to be different between two samples or sample sets in Mass Profiler, or between two or more sample sets via more sophisticated statistical analysis in Mass Profiler Professional. This scheme can be applied to any type of differential analysis, including applications in metabolomics, proteomics, environmental studies, food alteration, food origin, or forensics. Users can trigger the search directly from Mass Profiler or Mass Profiler Professional, and can review the results there.

Multiple workflows in MassHunter Qualitative Analysis

The majority of searches of AM(RT) databases are initiated from within MassHunter Qualitative Analysis. This software increases productivity by enabling compound-centric data analysis, which automatically finds compounds in complex LC/MS and LC/MS/MS analyses via proprietary algorithms. The software provides a number of workflows to find compounds and conduct searches of an AM(RT) database. (See **Figure 2**.) These flexible workflows, which are discussed in detail below, adapt the software for efficient analyses regardless of the application.

Screening for target compounds only

Two workflows are particularly effective when screening for target compounds such as pesticides. As shown in **Figure 2**, examples A and B, the workflows differ in the Find Compounds algorithm that is used. In both cases, the investigator starts with a list of target compounds with formulas in a database – either an Agilent or a custom Personal Compound Database. If the database contains non-target compounds in addition to target compounds, these workflows will search for both.

The first approach finds compounds using the Molecular Feature Extraction (MFE) algorithm. The user selects a setting that directs the software to filter the compound list by mass list, simply by pointing to the masses in the database. This limits the list of found compounds to only the ones of interest to the analyst. The software also reports the formula for each of the found target compounds as well as a score based on accurate mass match, isotopic ratios, and isotope spacing. Optionally, a full AM(RT) database search can be conducted, which gives access to the n best hits. This allows crosschecking whether some features might match to more than one compound.

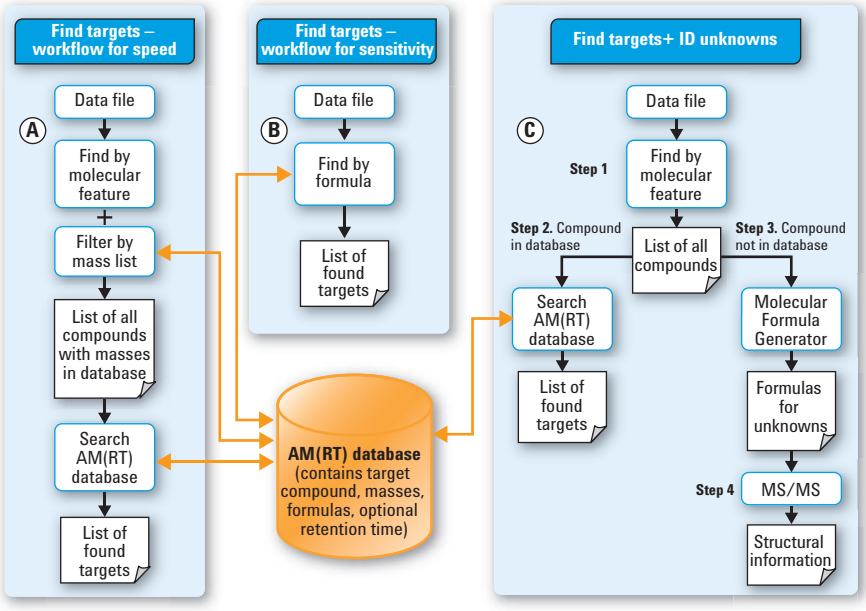

Figure 2. Compound-centric workflows, such as the examples shown here, can be automated and customized for any application.

The second approach uses the Find by Formula algorithm. The parameters for Find by Formula are set to search for the formulas of all the target compounds, which again is simply done by pointing to the respective accurate mass database. The software returns a list of those that were found.

The choice of workflow depends on the needs of the analyst. The first option (Find by MFE) is faster when searching a large database, while the second option (Find by Formula) provides greater sensitivity because it uses an extracted ion chromatogram (EIC) approach to find the target compounds. The Find by Formula option may require manual inspection of low-level results to eliminate false positives, but this can be done quickly with the compound-centric data navigation in the MassHunter Qualitative Analysis software. The analyst can pick either the Find by MFE approach or the Find by Formula approach, depending on whether it is more important to optimize speed or sensitivity.

Screening for both target and unknown compounds

For applications in the environmental area, as well as in forensics and toxicology, it is often desirable to both screen for target compounds and identify unknowns. Both types of analysis can be accomplished in a single run, using an automated workflow in MassHunter Qualitative Analysis. The process, shown in **Figure 2C**, is as follows:

1. Use Find Compounds by Molecular Feature to find all detectable compounds in the sample.
2. Perform an AM(RT) database search to find all target (and non-target) compounds.
3. Continue with Molecular Formula Generation (MFG) on the compounds that were not identified in the AM(RT) database (which at that point are unknowns). Another Agilent technical overview describes the functionality of the MFG.[3]
4. Export compound information for follow-up targeted MS/MS acquisition with an Agilent Q-TOF system. The accurate-mass MS/MS spectra are used for enhanced MFG and for correlation with proposed structures.

All of the steps above, except the correlation of the MS/MS spectra with proposed structures, can be automated. Therefore, this compound-centric workflow maximizes the amount of information extracted from a single analysis (see **Figure 3**), while minimizing the amount of time the investigator needs to spend on data processing.

MS/MS library search

The new LC/MS/MS accurate-mass library search capabilities, which are integrated into the new Personal Compound Database and Library (PCDL) software, provide even stronger confirmation for compound identification. This level of confirmation may be required for legal purposes (for example, in forensics and toxicology applications).

Compound List

Name	Cpd	RT	Base Peak	Height	Vol	Vol	Ion	Mass	Formula (MFG)	Mass (MF	Diff (MFG, pp	Score (MFG)	Formula (DB)	Mass (DB)	Diff (DB, ppm	Score (DB)	Hits (DB
Cymiazole	54	6.118	219.0954	428446	2139577	1.26	5	218.0881					C12 H14 N2 S	218.0878	-1.5	97.36	1
Chlorfluazuron	219	19.45	539.9706	101795	1189649	0.7	8	538.9633					C20 H9 Cl3 F5 N3 O3	538.963	-0.63	97.88	1
Carbofuranphenol	86	9.913	165.0911	58507	301767	0.18	3	164.0838					C10 H12 O2	164.0837	-0.52	99.17	1
Carbofuran, - 3 hydroxy	41	5.781	238.108	46190	417272	0.25	8	237.1006					C12 H15 N O4	237.1001	-2	95.43	1
Carbofuran	85	9.913	222.1128	241067	1172822	0.69	6	221.1055					C12 H15 N O3	221.1052	-1.47	84.14	1
Carbendazim	22	1.946	192.077	283383	1309753	0.77	5	191.0697					C9 H9 N3 O2	191.0695	-1.19	98.72	1
Carbaryl	89	10.427	202.0865	32623	188649	0.11	7	201.0792					C12 H11 N O2	201.079	-0.95	99.62	2
Bromuconazole	117	13.241	377.9592	71481	640404	0.38	8	374.9543					C13 H12 Br Cl2 N3 O	374.9541	-0.48	98.42	1
Bromuconazole	143	13.915	377.9594	51531	455539	0.27	8	374.9544					C13 H12 Br Cl2 N3 O	374.9541	-0.9	96.58	1
Boscalid	142	13.9	343.0405	47050	364275	0.21	6	342.0333					C18 H12 Cl2 N2 O	342.0327	-1.7	97.85	1
Azoxystrobin	135	13.688	404.1246	414838	2010805	1.18	5	403.1173					C22 H17 N3 O5	403.1168	-1.25	96.73	1
Acetamiprid	64	6.596	223.0747	91700	761953	0.45	15	222.0675					C10 H11 Cl N4	222.0672	-1.02	98.34	3
6-Benzylaminopurine	115	13.162	226.1077	168851	728454	0.43	3	225.1004					C12 H11 N5	225.1014	4.48	88.86	1
	3	0.745	216.9232	28861	102335	0.06	3	215.9159	C4 H7 Cl P2 S2	215.9153	-2.95	90.56					
	4	0.779	203.0531	488365	3116729	1.83	4	202.0458	C4 H6 N6 O4	202.0451	-3.79	74.05					
	5	0.782	365.1064	23659	204439	0.12	3	364.0991	C12 H23 N4 O3 P3	364.0983	-2.13	94.54					
	10	1.175	142.0086	77358	409173	0.24	6	123.9748	C3 H5 Cl O S	123.975	1.46	70.41					
	11	1.184	141.0772	65008	241925	0.14	2	140.0699	C5 H8 N4 O	140.0698	-0.95	87.44					
	12	1.3	142.9928	117644	688484	0.41	3	141.9855	C4 H4 N2 P2	141.985	-3.77	82.02					
	13	1.323	112.0759	38020	220322	0.13	2	111.0687	C6 H9 N O	111.0684	-2.3	85.58					
	14	1.324	113.0598	26345	207011	0.12	4	112.0525	C6 H8 O2	112.0524	-0.52	87.23					
	15	1.36	142.0863	25371	123127	0.07	2	141.079	C7 H11 N O2	141.079	-0.42	78.34					
	17	1.475	127.1119	32890	231855	0.14	3	126.1046	C8 H14 O	126.1045	-0.86	92.87					
	18	1.479	145.1225	25075	138951	0.08	3	144.1153	C8 H16 O2	144.115	-1.58	84.95					

Figure 3. Compound results table with database hits for target compounds, and MFG results on unknown compounds, from a single analysis.

Conclusion

The integration of Agilent Personal Compound Databases with the Agilent MassHunter Workstation software brings new levels of productivity to TOF-based target compound screening. The combination of Agilent MassHunter Workstation software, the Agilent Personal Compound Databases, and the PCDL software greatly increases the throughput and confidence of screening by providing:

- The ability to search from the MassHunter PCDL software, MassHunter Qualitative Analysis, or the Agilent software products for differential analysis (Mass Profiler and Mass Profiler Professional)
- Highly integrated workflows that can be customized to the needs of the application
- Sophisticated compound searching algorithms that remove the need for tedious manual data processing
- High confidence results, via scoring using accurate mass, isotope spacing, and isotope abundance in the MS domain, and the use of accurate mass MS/MS information
- Complete automation of the analysis, from spectral acquisition to customized Excel reports, for large batches of samples

These features increase the speed and reduce the cost of screening while providing accurate results.

References

1. F. Kuhlmann, P. Stone, and K. Imatani, "Increased Productivity for Target Compound Screening Using TOF-MS and Accurate-Mass Databases with Optional Retention Time, Part 1: Unique Capabilities of Agilent Software," Agilent publication number 5990-4828EN, **2009**.
2. J. Zweigenbaum, "Pesticide Personal Compound Database for Screening and Identification," Agilent publication number 5990-3976EN, **2009**.
3. E. Darland, D. McIntyre, D. Weil, F. Kuhlmann, and X. Li, "Superior Molecular Formula Generation from Accurate-Mass Data," Agilent publication number 5989-7409EN, **2008**.

Authors

Frank Kuhlmann is a Scientist and Product Manager for LC/MS Software, Pete Stone is an Applications Chemist, and Ken Imatani is a Product Manager for TOF and Q-TOF LC/MS systems, all at Agilent Technologies in Santa Clara, California, U.S.A.

www.agilent.com/chem/masshunter

This item is intended for Research Use Only. Not for use in diagnostic procedures.

Information, descriptions and specifications in this publication are subject to change without notice. Agilent Technologies shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance or use of this material.

© Agilent Technologies, Inc. 2009

Printed in the U.S.A. December 10, 2009
5990-4829EN

Agilent Technologies